RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/565,230_

Source: _IFWD_

Date Processed by STIC: _11/28/06_

ENTERED

CRF Errors Edited by the STIC Systems Branch

Serial Number: _10/565,230_

CRF Edit Date: _11/28/06_

Edited by: _____

____ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

____ Corrected the SEQ ID NO. Sequence numbers edited were:

_____

____ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's · edited:

_____

✓ Deleted: ___ invalid beginning/end-of-file text ; ___ page numbers

____ Inserted mandatory headings/numeric identifiers, specifically:

_____

____ Moved responses to same line as heading/numeric identifier, specifically:

_____

✓ Other: _Corrected <150>, <151> numeric identifiers_

_____

_____

IFWO

RAW SEQUENCE LISTING DATE: 11/28/2006
PATENT APPLICATION: US/10/565,230 TIME: 17:11:37

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF4\11282006\J565230.raw**

```
   3 <110> APPLICANT: TRANSGENE SA
   4       ERBS, Philippe
   6 <120> TITLE OF INVENTION: Polypeptide having an improved Cytosine deaminase
   7       activity
   9 <130> FILE REFERENCE: D21447
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/565,230
C--> 11 <141> CURRENT FILING DATE: 2006-01-20
  11 <150> PRIOR APPLICATION NUMBER: PCT/IB2004/002505
  12 <151> PRIOR FILING DATE: 2004-06-29
  14 <150> PRIOR APPLICATION NUMBER: US 60/508 274
  15 <151> PRIOR FILING DATE: 2003-10-06
  17 <150> PRIOR APPLICATION NUMBER: EP 03/360 087
  18 <151> PRIOR FILING DATE: 2003-07-21
  20 <160> NUMBER OF SEQ ID NOS: 2
  22 <170> SOFTWARE: PatentIn Ver. 2.1
  24 <210> SEQ ID NO: 1
  25 <211> LENGTH: 373
  26 <212> TYPE: PRT
  27 <213> ORGANISM: Artificial Sequence
  29 <220> FEATURE:
  30 <223> OTHER INFORMATION: Description of Artificial Sequence :Fusion
  31       protein having a CDase activity
  33 <300> PUBLICATION INFORMATION:
  35 <400> SEQUENCE: 1
  36 Met Val Thr Gly Gly Met Ala Ser Lys Trp Asp Gln Lys Gly Met Asp
  37   1               5                  10                  15
  39 Ile Ala Tyr Glu Glu Ala Ala Leu Gly Tyr Lys Glu Gly Gly Val Pro
  40                  20                  25                  30
  42 Ile Gly Gly Cys Leu Ile Asn Asn Lys Asp Gly Ser Val Leu Gly Arg
  43              35                  40                  45
  45 Gly His Asn Met Arg Phe Gln Lys Gly Ser Ala Thr Leu His Gly Glu
  46          50                  55                  60
  48 Ile Ser Thr Leu Glu Asn Cys Gly Arg Leu Glu Gly Lys Val Tyr Lys
  49   65                  70                  75                  80
  51 Asp Thr Thr Leu Tyr Thr Thr Leu Ser Pro Cys Asp Met Cys Thr Gly
  52                  85                  90                  95
  54 Ala Ile Ile Met Tyr Gly Ile Pro Arg Cys Val Val Gly Glu Asn Val
  55              100                 105                 110
  57 Asn Phe Lys Ser Lys Gly Glu Lys Tyr Leu Gln Thr Arg Gly His Glu
  58          115                 120                 125
  60 Val Val Val Val Asp Asp Glu Arg Cys Lys Lys Ile Met Lys Gln Phe
  61      130                 135                 140
  63 Ile Asp Glu Arg Pro Gln Asp Trp Phe Glu Asp Ile Gly Glu Ala Ser
```

RAW SEQUENCE LISTING   DATE: 11/28/2006
PATENT APPLICATION: US/10/565,230   TIME: 17:11:37

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF4\11282006\J565230.raw

```
64 145              150              155              160
66 Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn Gln Leu Leu
67              165          170          175
69 Gly Leu Tyr Thr Ile Ile Ser Asn Lys Asn Thr Thr Arg Pro Asp Phe
70              180          185          190
72 Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu Glu Gly Leu
73              195          200          205
75 Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp Thr Asn Glu
76       210          215          220
78 Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly Val Ser Ile
79 225              230          235              240
81 Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp Cys Cys Arg
82              245          250          255
84 Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu Glu Thr Ala
85              260 ·        265          270
87 Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile Ser Glu Arg
88       275          280          285
90 Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly Ser Ala Ile
91       290          295          300
93 Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro Glu Arg Ile
94 305              310          315              320
96 Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu Lys Tyr His
97              325          330          335
99 Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu Asp Arg Gly
100             340          345          350
102 Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp Phe Gly Asp
103        355          360          365      .
105 Arg Tyr Tyr Cys Val
106        370
109 <210> SEQ ID NO: 2
110 <211> LENGTH: 216
111 <212> TYPE: PRT
112 <213> ORGANISM: Saccharomyces cerevisiae
114 <400> SEQUENCE: 2
115 Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn
116 1            5            10           15
118 Gln Leu Leu Gly Leu Tyr Thr Ile Ile Ser Asn Lys Asn Thr Thr Arg
119             20           25           30
121 Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu
122        35           40           45
124 Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp
125    50           55           60
127 Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly
128 65           70           75           80
130 Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp
131             85           90           95
133 Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu
134        100          105          110
136 Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile
```

RAW SEQUENCE LISTING   DATE: 11/28/2006
PATENT APPLICATION: US/10/565,230   TIME: 17:11:37

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF4\11282006\J565230.raw

```
137           115                 120                 125
139 Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly
140       130                 135                 140
142 Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro
143 145                 150                 155                 160
145 Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu
146                 165           .     170                 175
148 Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu
149             180                 185                 190
151 Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
152         195                 200                 205
154 Phe Gly Asp Arg Tyr Tyr Cys Val
155     210                 215
```

VERIFICATION SUMMARY   DATE: 11/28/2006
PATENT APPLICATION: **US/10/565,230**   TIME: 17:11:38

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF4\11282006\J565230.raw**

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

Raw Sequence Listing before editing (for reference only)

RAW SEQUENCE LISTING   DATE: 11/24/2006
PATENT APPLICATION: US/10/565,230   TIME: 13:35:37

Input Set : **A:\PTO.DA.txt**
Output Set: **N:\CRF4\11242006\J565230.raw**

```
    3 <110> APPLICANT: TRANSGENE SA
    4        ERBS, Philippe
    6 <120> TITLE OF INVENTION: Polypeptide having an improved Cytosine deaminase
    7        activity
    9 <130> FILE REFERENCE: D21447
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/565,230
C--> 12 <141> CURRENT FILING DATE: 2006-01-20
   14 <150> PRIOR APPLICATION NUMBER: US 60/508 274
   15 <151> PRIOR FILING DATE: 2003-10-06
   17 <150> PRIOR APPLICATION NUMBER: EP 03/360 087
   18 <151> PRIOR FILING DATE: 2003-07-21
   20 <160> NUMBER OF SEQ ID NOS: 2
   22 <170> SOFTWARE: PatentIn Ver. 2.1
```

ERRORED SEQUENCES

p.2

```
  109 <210> SEQ ID NO: 2
  110 <211> LENGTH: 216
  111 <212> TYPE: PRT
  112 <213> ORGANISM: Saccharomyces cerevisiae
  114 <400> SEQUENCE: 2
  115 Met Ser Ser Glu Pro Phe Lys Asn Val Tyr Leu Leu Pro Gln Thr Asn
  116   1               5                  10                  15
  118 Gln Leu Leu Gly Leu Tyr Thr Ile Ile Ser Asn Lys Asn Thr Thr Arg
  119              20                  25                  30
  121 Pro Asp Phe Ile Phe Tyr Ser Asp Arg Ile Ile Arg Leu Leu Val Glu
  122          35                  40                  45
  124 Glu Gly Leu Asn His Leu Pro Val Gln Lys Gln Ile Val Glu Thr Asp
  125      50                  55                  60
  127 Thr Asn Glu Asn Phe Glu Gly Val Ser Phe Met Gly Lys Ile Cys Gly
  128  65                  70                  75                  80
  130 Val Ser Ile Val Arg Ala Gly Glu Ser Met Glu Gln Gly Leu Arg Asp
  131              85                  90                  95
  133 Cys Cys Arg Ser Val Arg Ile Gly Lys Ile Leu Ile Gln Arg Asp Glu
  134              100                 105                 110
  136 Glu Thr Ala Leu Pro Lys Leu Phe Tyr Glu Lys Leu Pro Glu Asp Ile
  137          115                 120                 125
  139 Ser Glu Arg Tyr Val Phe Leu Leu Asp Pro Met Leu Ala Thr Gly Gly
  140      130                 135                 140
  142 Ser Ala Ile Met Ala Thr Glu Val Leu Ile Lys Arg Gly Val Lys Pro
  143 145                 150                 155                 160
  145 Glu Arg Ile Tyr Phe Leu Asn Leu Ile Cys Ser Lys Glu Gly Ile Glu
```

```
     146                   165                170                175
     148 Lys Tyr His Ala Ala Phe Pro Glu Val Arg Ile Val Thr Gly Ala Leu
     149             180                185                190
     151 Asp Arg Gly Leu Asp Glu Asn Lys Tyr Leu Val Pro Gly Leu Gly Asp
     152         195                200                205
     154 Phe Gly Asp Arg Tyr Tyr Cys Val
     155       210             215
E--> 163 (1
```

VERIFICATION SUMMARY   DATE: 11/24/2006
PATENT APPLICATION: **US/10/565,230**   TIME: 13:35:38

Input Set : **A:\PTO.DA.txt**
Output Set: **N:\CRF4\11242006\J565230.raw**

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:163 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:2